Vectores Autoregresivos - VAR

Leonardo Sánchez Aragón

August 6, 2020

Introducción

- Cuando no tenemos la certeza de que una variable es exógena, una extensión del ande función de trasnferencia es VAR.
- Consideremos el caso de dos variables que serán endógenas: la secuencia de valores $\{y_t\}$ son efectados por valores pasados de si mismo y los valores actuales y pasados de $\{z_t\}$. De igual fomar sucede con la secuencia $\{z_t\}$

$$y_t = b_{10} - b_{12}z_t + \gamma_{11}y_{t-1} + \gamma_{12}z_{t-1} + \varepsilon_{yt}$$
$$z_t = b_{20} - b_{21}y_t + \gamma_{21}y_{t-1} + \gamma_{22}z_{t-1} + \varepsilon_{zt}$$

- Este es un Vector Autoregresivo de orden 1 - VAR(1). Este sistema se conoce como VAR Estructural.
- Se asume que: (i) y_t y z_t son estacionarias; (ii) ε_{yt} y ε_{zt} son ruido blanco con desviaciones σ_y y σ_z; y (iii) ε_{yt} y ε_{zt} están no correlacionados.

Introducción

- El coeficiente b_{12} es el efecto contemporáneo de z_t sobre y_t, y γ_{12} es el efecto de z_{t-1} sobre y_t.
- Si $b_{21} \neq 0$, entonces ε_{y_t} tiene un efecto indirecto contemparaneo sobre z_t, y si $b_{12} \neq 0$, entonces ε_{z_t} tiene un efecto indirecto contemparaneo sobre y_t.
- Las ecuaciones previas no se pueden estimar usando MCO ya que y_t tiene un efecto contemporaneo sobre z_t y esta sobre y_t. Las estimaciones de MCO sufrirían un sesgo de ecuación simultánea ya que los regresores y los términos de error estarían correlacionados.
- Reordenando el sistema, se pudiera estimar. Matricialmente tenemos,

$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

$$Bx_t = \Gamma_0 + \Gamma_1 x_{t-1} + \varepsilon_t$$

Introducción

- Donde las matrices de $Bx_t = \Gamma_0 + \Gamma_1 x_{t-1} + \varepsilon_t$, son

$$B = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix}, x_t = \begin{bmatrix} y_t \\ z_t \end{bmatrix}, \Gamma_0 = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix}, \Gamma_1 = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}, \varepsilon_t = \begin{bmatrix} \varepsilon_{y_t} \\ \varepsilon_{z_t} \end{bmatrix}$$

- Premultiplicando por B^{-1}, tenemos

$$x_t = A_0 + A_1 x_{t-1} + e_t$$

Donde $A_0 = B^{-1}\Gamma_0$, $A_1 = B^{-1}\Gamma_1$ y $e_t = B^{-1}\varepsilon_t$

- Se define a_{i0} como el i-esimo elemento del vector A_0, a_{ij} es el elmento de la fila i y columna j de la matriz A_1, e_{it} es el elemento i del vector e_t.

Introducción

- Entonces, el sistema se puede reescribir como
  $y_t = a_{10} + a_{11} y_{t-1} + a_{12} z_{t-1} + e_{1t}$
  $z_t = a_{20} + a_{21} y_{t-1} + a_{22} z_{t-1} + e_{2t}$
- Esta forma de representar el VAR se denomina forma estandar.
- Cabe destacar que los errores $e_t = B^{-1} \varepsilon_t$
  $e_{1t} = \left(\varepsilon_{y_t} - b_{12} \varepsilon_{z_t}\right) / \left(1 - b_{12} b_{21}\right)$
  $e_{2t} = \left(\varepsilon_{z_t} - b_{21} \varepsilon_{y_t}\right) / \left(1 - b_{12} b_{21}\right)$
- Dado que ε_{z_t} y ε_{y_t} son ruidos blancos, entonces e_{1t} y e_{2t} tiene media cero, varianza constante y son individualmente no correlacionados.

Introducción

- El valor esperado es:

$$Ee_{1t} = E\left(\varepsilon_{yt} - b_{12}\varepsilon_{zt}\right) / \left(1 - b_{12}b_{21}\right) = 0$$

- La varianza es:

$$\begin{aligned} Ee_{1t}^2 &= E\left[\left(\varepsilon_{yt} - b_{12}\varepsilon_{zt}\right) / \left(1 - b_{12}b_{21}\right)\right]^2 \\ &= \left(\sigma_y^2 + b_{12}^2\sigma_z^2\right) / \left(1 - b_{12}b_{21}\right)^2 \end{aligned}$$

- La correlación es

$$\begin{aligned} Ee_{1t}e_{2t} &= E\left[\left(\varepsilon_{yt} - b_{12}\varepsilon_{zt}\right)\left(\varepsilon_{zt} - b_{21}\varepsilon_{yt}\right)\right] / \left(1 - b_{12}b_{21}\right)^2 \\ &= -\left(b_{21}\sigma_y^2 + b_{12}\sigma_z^2\right) / \left(1 - b_{12}b_{21}\right)^2 \end{aligned}$$

- Si $b_{12} = b_{21} = 0$ los errores serán no correlacionados.

Introducción

- Es útil escribir la matriz de varianza y covarianza

$$\Sigma = \begin{bmatrix} \mathrm{var}\left(e_{1t}\right) & \mathrm{cov}\left(e_{1t}, e_{2t}\right) \\ \mathrm{cov}\left(e_{1t}, e_{2t}\right) & \mathrm{var}\left(e_{2t}\right) \end{bmatrix}$$

- En forma compacta tenemos

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}$$

Estacionariedad y Estabilidad

- La ecuación $x_t = A_0 + A_1 x_{t-1} + e_t$, puede resolverse de manera recursiva como

$$x_t = (I + A_1 + ... + A_1^n)A_0 + \sum_{i=0}^{n} A_1^i e_{t-i} + A_1^{n+1} x_{t-n-1}$$

- Condición de estabiliadad: Se requiere que la expresión A_1^n se desvanezca a medida que n tiende al infinito.
- Dicho de otra forma: se requiere que las raíces de $(1 - a_{11}L)(1 - a_{22}L) - (a_{12}a_{21}L^2)$ estén fuera del círculo unitario.

Estacionariedad y Estabilidad

- Si la condición de estabilidad se cumple la solución particular para x_t es

$$x_t = \mu + \sum_{i=0}^{\infty} A_1^i e_{t-i}$$

- Donde $\mu = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix}$

$$\begin{aligned} \bar{y} &= \left[a_{10}\left(1-a_{22}\right)+a_{12}a_{20}\right]/\Delta \\ \bar{z} &= \left[a_{20}\left(1-a_{11}\right)+a_{21}a_{10}\right]/\Delta \\ \Delta &= \left(1-a_{11}\right)\left(1-a_{22}\right)-a_{12}a_{21} \end{aligned}$$

Estacionariedad y Estabilidad

- Si tomamos el valor esperado de $x_t = \mu + \sum_{i=0}^{\infty} A_1^i e_{t-i}$, el valor esperado incondicional es μ
- La matriz de varianza y covarianza es

$$
\begin{aligned}
E\left(x_t - \mu\right)^2 &= E\left[\sum_{i=0}^{\infty} A_1^i e_{t-i}\right]^2 \\
&= \left(I + A_1^2 + A_1^4 + A_1^6 + \cdots\right)\Sigma \\
&= \left[I - A_1^2\right]^{-1}\Sigma
\end{aligned}
$$

Donde

$$
Ee_t^2 = \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} \begin{bmatrix} e_{1t} & e_{2t} \end{bmatrix} = \Sigma
$$

Estimación

- Considere el siguiente modelo

$$x_t = A_0 + A_1 x_{t-1} + A_2 x_{t-2} + \ldots + A_p x_{t-p}. + e_t$$

- Se asume que las variables incorporadas en el VAR responde a un modelo económico (Sims, 1980).
- Este sistema de ecuaciones se estima por Minimo Cuadrados Generalizados.
- Los coeficientes estimados son idénticos a los obtenidos al estimar ecuación por ecuación, ya que todas las regresiones tienen variables idénticas en el lado derecho (Zellner, 1962), pero las desvaciones standard pueden diferir si las ecuaciones están relacionadas
- Otro método de estimacion es usar Máxima Verosimilitud, pero se asume que se conoce la distribución de los errores del sistema.

Estimación

- La matriz A_0 contiene n parámetros, y cada matriz A_i contiene n^2 parámetros; por lo tanto, se necesitan estimar los coeficientes $n + pn^2$.
- El VAR estará sobreparamizado porque muchas de estas estimaciones de coeficientes serán insignificantes.
- Solución: imponer restricciones sobre los parámetros del VAR estructural. Pero la imposición incorrecta de restricciones cero puede desperdiciar información importante. Además, es probable que los regresores sean altamente colineales, por lo que las pruebas t de coeficientes individuales no son guías confiables para reducir el modelo.

Identificación

- Cuando estima el modelo en forma estandar hay 6 parámetros estimados $(a_{10}, a_{20}, a_{11}, a_{12}, a_{21}, a_{22})$ y los valores de $var(e_{1t})$, $var(e_{2t})$, $cov(e_{1t}, e_{2t})$
- El VAR estructural tiene 10 paráametros: $(b_{10}, b_{20}, \gamma_{11}, \gamma_{12}, \gamma_{21}, \gamma_{22})$, los efectos contemporáneos $(b_{12}$ y $b_{21})$ y los valores de σ_y^2, σ_z^2
- Para recuperar los parámetros originales debemos imponer una restricción al modelo estructural.

Identificación

- Una poropuesta es la sugerida por Sims 1980. Si estamos dispuesto a eliminar el impacto de y_t sobre z_t (es decir, $b_{21} = 0$), el sistema se hace asimétrico. Esto permite recuperar todos los coeficientes.

$$\begin{aligned} e_{1t} &= \varepsilon_{yt} - b_{12}\varepsilon_{zt} \\ e_{2t} &= \varepsilon_{zt} \end{aligned}$$

Tal que

$$\begin{aligned} \operatorname{var}(e_1) &= \sigma_y^2 + b_{12}^2\sigma_z^2 \\ \operatorname{var}(e_2) &= \sigma_z^2 \\ \operatorname{cov}(e_1, e_2) &= -b_{12}\sigma_z^2 \end{aligned}$$

Identificación

- Viendo la restricción desde otra perspectiva, tenemos.

$$\underbrace{\begin{bmatrix} 1 & b_{12} \\ 0 & 1 \end{bmatrix}}_{B} \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

- Pre multiplicando por B^{-1}, tenemos

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} b_{10} - b_{12}b_{20} \\ b_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} - b_{12}\gamma_{21} & \gamma_{12} - b_{12}\gamma_{22} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} - b_{12}\varepsilon_{zt} \\ \varepsilon_{zt} \end{bmatrix}$$

$$\begin{aligned} y_t &= a_{10} + a_{11}y_{t-1} + a_{12}z_{t-1} + e_{1t} \\ z_t &= a_{20} + a_{21}y_{t-1} + a_{22}z_{t-1} + e_{2t} \end{aligned}$$

Identificación

- El supuesto $b_{21} = 0$ significa que y_t no tiene un efecto contemporáneo sobre z_t. La restricción se manifiesta de tal manera que los shocks ε_{y_t} y ε_{z_t} afectan el valor contemporáneo de y_t, pero solo los choques ε_{z_t}afectan el valor contemporáneo de z_t. Los valores observados de e_{2t} se atribuyen completamente a shock puros a la secuencia z_t. Esta descomposición triangular de los residuos se llama **descomposición de Choleski**.
- De hecho, el resultado es bastante general. En un VAR de n variables, B es una matriz $n \times n$ ya que hay n residuos de regresión y n shochs estructurales. La identificación exacta requiere que se coloquen $(n^2 - n)/2$ restrictions entre los residuos de regresión. Dado que la descomposición de Choleski es triangular, se fuerza exactamente a $(n^2 - n)/2$ valores de la matriz B a ser igual a cero.

Identificación

- Hay un problema si las variables en un VAR no son estacionarias? Sims (1980) y Sims, Stock y Watson (1990) recomendaron no diferenciar, incluso si las variables contienen una raíz unitaria.
- Argumentan que el objetivo de un análisis VAR es determinar las interrelaciones entre las variables, no determinar las estimaciones de los parámetros.
- El argumento principal en contra de la diferenciación es que "desecha" información sobre los comovimientos en los datos (como la posibilidad de relaciones de cointegración).
- Del mismo modo, se argumenta que no es necesario detrend los datos. En un VAR, una variable de tendencia estará bien aproximada por una raíz unitaria más drift. Sin embargo, la opinión mayoritaria es que la forma de las variables en el VAR debe imitar el verdadero proceso de generación de datos.

Selección de rezagos

- Cuántos rezagos incluir en el estimación?
- Un enfoque es estimar muchos modelos VAR con diferentes números de rezagos y comparar su ajuste.
- La literatura utiliza un conjunto de criterios de selección: el criterio de información de Akaike (AIC), el criterio de información bayesiano de Schwarz (SBIC) y el criterio de información de Hannan y Quinn (HQIC).

$$\mathrm{AIC} = T \ln |\Sigma| + 2N$$
$$\mathrm{SBC} = T \ln |\Sigma| + N \ln(T)$$

Donde $|\Sigma|$ es el determinante de la matriz de varianza y covarianza de los residuos, y N es el número total de parámetros estimados en todas las ecuaciones. Agregar regresores adicionales reducirá $\ln |\Sigma|$ a expensas de aumentar N.

Función Impulso Respuesta

- El modelo VAR tiene un representaciobn VMA.

$$x_t = \mu + \sum_{i=0}^{\infty} A_1^i e_{t-i}$$

- La representacion VMA permite expresar las vairanles y_t y z_t en terminos de los errores (e_{1t}, e_{2t}) presents y pasados
- Esta representacion permite delinear a traves del tiempo el efecto de los shock sobre las variabels presentes en el VAR

Función Impulso Respuesta

- A partir de la representación VMA, tenemos

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} e_{1t-i} \\ e_{2t-i} \end{bmatrix}$$

- Recordando que $e_t = B^{-1}\varepsilon_t$, es decir,

$$\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \frac{1}{1 - b_{12}b_{21}} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

- Podemos unir las expresiones previias y obtener

$$\begin{bmatrix} y_t \\ z_i \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \frac{1}{1 - b_{12}b_{21}} \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{yt-i} \\ \varepsilon_{zt-i} \end{bmatrix}$$

Función Impulso Respuesta

- La expresión previa puede ser resumida mediante la simplificación de la matriz 2×2

$$\phi_i = \frac{A_1^i}{1 - b_{12} b_{21}} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix}$$

- Esto permite reexpresar en modelo en función de los valores de ε_t

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \bar{y} \\ \bar{z} \end{bmatrix} + \sum_{i=0}^{\infty} \begin{bmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{bmatrix} \begin{bmatrix} \varepsilon_{yt-i} \\ \varepsilon_{zt-i} \end{bmatrix}$$

$$x_t = \mu + \sum_{i=0}^{\infty} \phi_i \varepsilon_{t-i}$$

Función Impulso Respuesta

- Los elementos $\phi_{jk}(0)$ son multiplicadores de impacto. Si $\phi_{12}(0)$ es el instantáneo impacto de un cambio de una unidad de ε_{z_t} en y_t.
- Los elementos $\phi_{11}(1)(1)$ y$\phi_{12}(1)$ son las respuestas luego de un período de los cambios unitarios en $\varepsilon_{z_{t-1}}$ y$\varepsilon_{y_{t-1}}$ sobre y_t, respectivamente.
- Los efectos acumulados de los impulsos unitarios en $\varepsilon_{z_{t-1}}$ y $\varepsilon_{y_{t-1}}$ se pueden obtener mediante la suma adecuada de los coeficientes de las funciones de impulso respuesta.
- Por ejemplo, después de n períodos, el efecto de $\varepsilon_{z_{t-1}}$ en el valor de y_{t+n} es $\phi_{12}(n)$ Así,después de n períodos, la suma acumulada de los efectos de $\varepsilon_{z_{t-1}}$ en la secuencia y_t es

$$\sum_{i=0}^{n} \phi_{12}(i)$$

Función Impulso Respuesta

- Si z_t en y_t son estacionarias, los valores de $\phi_{jk}(0)$ convergen a cero.
- Los cuatro conjuntos de coeficientes $\phi_{11}(i)$, $\phi_{12}(i)$, $\phi_{21}(i)$, $\phi_{22}(i)$ denominan funciones impulso respuesta.
- Si podemos recuperar los parametros del VAR estructural, podemos obtener los errores estructurales ε_{z_t} y ε_{y_t}, y con ellos los coeficientes $\phi_{jk}(i)$
- La Descomposición de Choleski permite la identificación de los parámetros estructurales: y_t no tiene un efecto contemporáneo sobre z_t. Formalmente, esta restricción se representa estableciendo $b_{21} = 0$.

Función Impulso Respuesta

- Los términos de error se pueden descomponer de la siguiente manera:

$$e_{1t} = \varepsilon_{yt} - b_{12}\varepsilon_{zt}$$
$$e_{2t} = \varepsilon_{zt}$$

- Se dice que este un "orden" de las variables. Una shock ε_{z_t} afecta directamente a e_{1t} y e_{2t}, pero una descarga ε_{y_t}no afecta a e_{2t}. Por lo tanto, se dice que z_t es "causa previamente" a y_t
- La Descomposición de Choleski restringe el sistema: un shock ε_{y_t} no tiene efecto directo sobre z_t, pero hay un efecto indirecto en que los valores rezagados de y_t afectan el valor contemporáneo de zt.
- La Descomposición fuerza una asimetría potencialmente importante en el sistema ya que un shcok ε_{z_t} tiene efectos contemporáneos en yt y zt.